



Consequences of Finite Numerical Precision in Iterated Maps

J. A. MacLachlan, H. Del Manzano*

Fermi National Accelerator Laboratory,[†] Box 500, Batavia IL 60510

October 2, 2001

Abstract

Numerical models of beam behavior in cyclic accelerators are typically obtained by iterating a map which transforms the phase space coordinates of representative particles to their coordinates one beam turn later. For such purposes as dynamic aperture determination or simulation of extended processes, many iterations, 10^7 or more, may be used. Therefore, it is important to establish that limited numerical precision does not result in unacceptable cumulative error. Adiabatic invariants have good stability with respect to numerical noise under iteration of a symplectic map. If the map is time independent, the effects of roundoff on the map itself are coherent; a symplectic map offers no protection from this sort of error. In diagnosing apparent cumulative error, it is very important to look at results for precisely matched initial distributions; otherwise numerical error can be confounded with the slow evolution of the non-stationary distribution. A numerical evaluation of the Jacobian determinant gives a useful quantitative measure of the precision in the numerical representation of a symplectic map. This paper specializes to a one-dimensional longitudinal phase space map for synchrotron motion. However, the approach is general enough to be applicable, at least in part, to other maps.

Introduction

The general idea of studying the evolution of a beam particle distribution in a cyclic accelerator by iterating single-particle, single-turn maps is ubiquitous in beam physics. An example of such a map is a transfer matrix produced by a matrix multiplication code like MAD[1] or TRANSPORT.[2] In these particular instances the component matrices are not symplectic above first order, and emittance of a distribution within a region of stable motion may grow or sometimes shrink.[3] The emittance change results from defects in the map and would appear even with ideal numerical precision. A symplectic map is one that represents an exact solution for some Hamiltonian; it preserves all of the Poincare invariants. A map is symplectic if it satisfies the condition

$$\mathcal{J}^T S \mathcal{J} = S \quad , \quad (1)$$

where S is the so-called symplectic unit matrix

$$\begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \quad (2)$$

and \mathcal{J} is the Jacobian matrix, the matrix whose determinant J , called the Jacobian, is employed in this note. An explicit expression for \mathcal{J} is evident from eq. 5 giving J for a particular map. Symplectic maps can be produced by an inherently symplectic technique like Lie transforms, for example, or they can be, as it is said, symplectified by adding the necessary terms. However, it is also possible that the effect of truncation error in evaluating the map might destroy the symplectic character at some level. This is subject to numerical test; the numerical evaluation of the Jacobian determinant at a precision greater than the precision used in the mapping can expose discrepancies in the phase space volume preservation — systematic or stochastic, general or localized in a region of phase space. For a one-dimensional map, area preservation and symplecticity are equivalent. For maps of higher dimension, the Jacobian test remains a useful diagnostic, but the symplecticity condition eq. 1 is the full test for symplecticity.

It has been shown by Forest[4] that a map in which the potential is applied in impulses interspersed with drifts is symplectic regardless of how faithful it is to the dynamics of the system in other respects. It constitutes a so-called kick integrator for the system's equations of motion. The equations for synchrotron motion with the rf potential

*Permanent address Dept. of Electrical Engineering, University of Puerto Rico, Mayaguez Campus, Mayaguez, P. R. 00681

[†]Work supported by the U.S. Department of Energy under contract No. DE-AC02-76CH03000.

appearing across an isolated gap fall naturally into this pattern without approximation:[5]

$$\begin{aligned}\varphi_{i,n} &= \frac{\omega_{s,n}}{\omega_{s,n-1}}\varphi_{i,n-1} + 2\pi h(S_{i,n} - 1) \\ \varepsilon_{i,n} &= \varepsilon_{i,n-1} + eV(\varphi_{i,n} + \phi_{s,n}) - eV(\phi_{s,n}) ,\end{aligned}\tag{3}$$

where φ is a phase difference between a particle and the synchronous phase ϕ_s , likewise ε is the energy difference between a particle and the synchronous energy E_s ; i labels particles, and n labels turns. All symbols are defined in Table I.

The φ, ε variables are not canonically conjugate, but it is simple to rewrite the map with the substitution $e_{i,n} = \varepsilon_{i,n}/\omega_{i,n}$ as the map \mathbf{M}

$$\begin{aligned}\varphi_{i,n} &= \frac{\omega_{s,n}}{\omega_{s,n-1}}\varphi_{i,n-1} + 2\pi h(S_{i,n} - 1) \\ e_{i,n} &= \frac{\omega_{s,n-1}}{\omega_{s,n}}e_{i,n-1} + \frac{e}{\omega_{s,n}}[V(\varphi_{i,n} + \phi_{s,n}) - V(\phi_{s,n})] .\end{aligned}\tag{4}$$

The Jacobian is

$$J(\mathbf{M}) = \left| \frac{\partial(\varphi_{i,n}, e_{i,n})}{\partial(\varphi_{i,n-1}, e_{i,n-1})} \right| = \left| \begin{array}{cc} \frac{\omega_{s,n}}{\omega_{s,n-1}} & \frac{e}{\omega_{s,n}}eV' \\ 2\pi h \frac{\partial S_{i,n}}{\partial e_{i,n-1}} & \frac{\omega_{s,n-1}}{\omega_{s,n}} + \frac{e}{\omega_{s,n}}V'2\pi h \frac{\partial S_{i,n}}{\partial e_{i,n-1}} \end{array} \right| \equiv 1\tag{5}$$

However, the coordinates are subject to numerical noise, *i. e.* rounding error, for both the kick and the drift in the map. Therefore, there is a continuous stochastic contribution to the trajectories which can be interpreted as an extremely small fluctuating force. One approach to estimating the effects of truncation is to estimate the Fourier spectrum of this force. For one dimensional motion the action is an adiabatic invariant proportional to the emittance. Therefore, one can anticipate that emittance growth will be strongly inhibited compared to fluctuation of the phase of motion along trajectories.[6] However, if the noise spectrum were to have significant strength at frequencies as low as twice the synchrotron frequency, then emittance growth would certainly occur. The preservation of emittance is not directly explained by the adiabatic theorem, because the high frequency components apply force on a time scale *short* with respect to the synchrotron period and yet have little effect on the adiabatic invariant.

Coherent Excitation from Truncation Error

In a tracking model, the map may be static or it may be time dependent. If the map changes, the error in evaluating the map at different times will be an additional random perturbation. However, if the same identical map is iterated many times, whatever error is made in its numerical representation is fixed and makes a coherent error in the macro-particle trajectories.

This paper arises from studies of the source of apparent numerical instability with the map eq. 3. Fig. 1 shows a phase space trajectory taken over 10^6 iterations for the centroid of a compact distribution of four particles initially centered on the stable fixed point. The parameters are fixed throughout the calculation, providing a non-accelerating bucket above the synchrotron transition energy. The coordinate axes in the figure are MeV on the ordinate and rf phase divided by the harmonic number on the abscissa. The parameters for the map are those of the Tevatron at injection energy (Table II), and the four test particles are at $\varphi = 180^\circ \pm 1^\circ$, $\varepsilon = \pm 1$ MeV. In this period of time the centroid trajectory has grown by about $2 \cdot 10^{-10}$ of the bucket width, that is about $2 \cdot 10^{-16}$ /turn. Note that there is an underlying dipole oscillation with amplitude about 10^{-9} of the bucket width; it is not spurious. It results from the dependence of phase-slip on energy. The particles below the synchronous energy slip less than those above, because the energies are above transition. Since the distribution is symmetric about the stable fixed point but not matched to the Hamiltonian flow, the macro-particle's centroid, which is started at (0,0), averages below zero in energy.

The centroid trajectory shown in Fig. 2 results from 10^6 iterations with an approximate map that employs the same slip factor for all macro-particles. In this figure the final trajectory extends only about $3 \cdot 10^{-16}$ of a bucket width. Thus, the average increment to the centroid is approximately $3 \cdot 10^{-22}$ /turn. If one looks at just 10^4 turns as in Fig. 3, the average systematic increment is larger but of the same order. However, on this scale one can see also the stochastic nature of the centroid motion; the centroid trajectory crosses itself many times. Because the major distinction of the simpler map is removing differences in slip velocity above and below the synchronous energy, it is a reasonable inference that much of the drift of the centroid recorded in Fig. 1 arises from the fact that the little four-particle distribution is not quite matched to the original map, *i. e.*, is not a stationary distribution. The outward spiral in Fig. 2 appears to be approximately linear; it does not exhibit exponential growth like a collective beam instability. Indeed there is no collective mechanism; only single particle motion in a nearly linear potential is involved.

After removing the effect of the small mismatch in the initial distribution by using the simplified map, there is still a small systematic outward spiral of centroid, a coherent change of the mean energy of the distribution. When the map is used below transition for 10^7 turns, the centroid plot appears as shown in Fig. 4. Not only is the centroid motion apparently fully stochastic in this case, but also the region of wander is two orders of magnitude smaller. Given that the same map is used above and below transition, attention is naturally drawn to the constant term $eV(\phi_{s,n})$ in the energy kick; for an ordinary rf waveform, this term is proportional to $\sin(0)$ below transition and to $\sin(\pi)$ above transition. A value for the sine differing from zero by 10^{-19} could introduce about this amount of energy/turn.

These phenomena may be called instability of numerical origin, but they are not of the stochastic type often discussed in a numerical analysis context. For the systematic energy increment there is a question whether the sine function or its argument is the culprit. The sine was absolved by evaluating it in small steps either side of π . In the following discussion numbers are described as double precision, meaning that their computer representation is eight byte floating point, or quadruple precision, meaning that the computer representation is sixteen eight-bit bytes. When π was evaluated in quadruple precision and the argument of the double precision sine was taken in steps of rational binary fractional increments of π in the range $(1 \pm 10^{-5})\pi$, the plot of the sine was apparently smooth and linear, missing zero at argument π by 10^{-51} . Furthermore, $\sin \varphi - \pi + \varphi$ produces a smooth cubic around the zero in this range, as shown in Fig. 5. These results show that the error in the sine function is not set by the scale of the maximum amplitude but rather scales with the value of the function. Thus, the observed energy increment must result from the truncation of the value of the argument π to double precision, *i. e.*, possibly an error of as much as $\sim 10^{-14}$. One may reasonably question whether such a small energy input is of practical importance. The usual answer would, of course, be no, but if one notices the phenomenon without knowing the cause, it may raise the question of the fundamental correctness of the formulation. The following examination of effects of limited numerical precision has the object of helping to validate map-based numerical models. Incidentally, in this special case, should the numerical error be a problem, there is a simple work-around; the plus sign in front of the second term on the right-hand side of the first equation in the map can be changed to a minus sign above the transition energy. Then the stable phase will remain at less than $\pi/2$ but the direction of phase flow will reverse as it should.

Numerical Test of Area Preservation

The map in eq. 3 is supposed to represent the equations of motion for a conservative process. Therefore, the phase space area enclosed by a trajectory should be conserved when parameters change slowly, and the map should be strictly area preserving. The discussion above illustrates that one can easily misjudge the cumulative error arising from the map by tracking an unmatched test distribution. Getting a highly matched initial distribution is not entirely simple, although, with care good approximations can be made. It is desirable to have a check of the map not depending on the properties of a test distribution. It is possible to verify that a map is not only area preserving in principle but also in practice by evaluating the Jacobian determinant numerically from coordinate differences. Additional calculations with the coordinates must be made at higher precision than used in the map itself to permit unambiguous association of error in the Jacobian with errors in the map alone. One can evaluate each of the four terms in the Ja-

cobian determinant eq. 5 analytically. This provides a check on the precision of the numerical evaluation in this case and helps to understand the practical tradeoff between precision of the difference approximation to the derivatives depending on small separation of coordinates and the loss of significant figures, which is reduced by larger separation of coordinates. The error in approximating the derivative might in principle be reduced by mapping more points and using a higher order formula. Given that the mapped points are not on a uniform grid, calculation of the derivatives above first order differences would be tedious.

In Fig. 6 we give a distribution of the Jacobian - 1 for eq. 3 evaluated numerically for a specific case which is, as before, like the Tevatron at injection. The evaluations are made on a 25×25 grid spaced at 10° intervals in rf phase and 10 MeV in energy. The grid is centered on the synchronous phase of 180° and contains some points outside of the bucket. The cell size for evaluation of the derivatives is $0.2^\circ \times 0.2$ MeV for this case. The longest bar in the plot is $1.1 \cdot 10^{-10}$. This choice of cell size is purely arbitrary, providing an upper limit on the difference. The limit can be optimized by seeking the minimum of the maximum difference as a function of cell size. Two more cases are shown: Fig. 7, with a longest bar of $3.3 \cdot 10^{-12}$, gives the distribution for cells of 0.02×0.02 and Fig. 8, with a longest bar of $1.1 \cdot 10^{-11}$, for cells of 0.005×0.005 . In Fig. 9, the maximum absolute value for the difference J - 1 is plotted as a function of the length of the cell edges. For the particular parameters of the example, the minimum value for the error in the Jacobian is obtained with cell dimensions of about $0.02^\circ \times 0.02$ MeV.

Even in Fig. 8, where the cancellation in evaluating the derivatives in the Jacobian appears dominant, the vestiges of the same systematic distribution are still apparent. Note that J - 1 is independent of E, and changes sign at $\pm\pi/2$; it appears to be essentially proportional to $\cos\varphi$. The systematic distribution almost surely arises from the map, with which it has a shared symmetry, rather than from the limitation to first order differences for the derivatives, which one would expect to show up in approximately comparable degree in both energy and phase differences. Given the inherent cancellation in the difference equations, an error of a few parts in 10^{12} in the Jacobian derivative is very satisfactory. Its systematic character looks consistent with the map.

Constancy of Adiabatic Invariants Under Stochastic Perturbation

Quantities which are adiabatic invariants are more stable than other functions of the coordinates under iteration with a map representing the equations of motion. This is familiar to many, but the underlying reason for the insensitivity of these quantities to rapidly fluctuating perturbation is not elementary. The discussion on pp 291 – 297 of Arnold’s *Mathematical Methods of Classical Mechanics*[6] provides the basis for understanding but does not explicitly state the result. In the present context of numerical tests it is interesting to compare a plot of longitudinal emittance vs. time with a plot of some other function of the coordinates like, for example, centroid position. This particular comparison will be good only if there is no comparable or greater coherent motion of the centroid.

Fig. 10 shows the scaled rms emittance minus 0.5 eVs for a nominal 0.5 eVs closed contour during 10^6 iterations, where the scaling sets the initial value to precisely 0.5. The parameters are again those of Table II for the above transition condition. It is constant to $\pm 10^{-5}$ with no change in mean value. This is an upper limit on the true fluctuation of the emittance, because the rms is not in principle a determination of a phase space area. The centroid locus, however, has a spread of a few 10^{-4} of the contour radius. Perhaps one might predict somewhat better constancy of the emittance on the basis of the other results on precision; however, it is at least an order of magnitude less variable than the mean and constant enough for any practical purpose. An analysis which avoids using the rms emittance might lower the limit for numerical fluctuation of the emittance.

Summary

The observation of coherent growth of synchrotron oscillation in tracking for a non-accelerated bunch has led to a study of effects of numerical error in extended simulations using iterated maps. The source of the growing oscillation was identified as the truncation error in the synchronous phase of π radians. Because this error was constant for the

entire calculation, its effect was cumulative and coherent. The effects of fluctuating truncation and cancellation error were also considered. The difficulty in distinguishing emittance growth resulting from the evolution of an imperfectly matched initial distribution from that arising from numerical error in the map has led to development of a direct, non-iterative test of the map. The numerical stability of adiabatic invariants compared to other functions of the coordinates was illustrated.

Acknowledgments

C. Y. Tan of Fermilab was the first to report the growth of small-scale synchrotron oscillation at fixed energy, like that shown in Fig. 1. We are indebted to Nadezhda Shemyakina of the State University of Novosibirsk and Leo Michelotti of Fermilab for the data visualization software which produced Figs. 5 – 7. HDM wishes to thank the SIST committee at Fermilab for their hospitality and giving him the opportunity to join this study.

References

- [1] H. Grote and F. C. Iselin, “MAD — Methodical Accelerator Design”, CERN Program Library, copyright CERN (1990)
- [2] D. C. Carey, *et al.*, “Third Order Transport with MAD Input”, SLAC Report 530
- [3] Norman Gelfand, “The Effect of Dividing the Magnets on the Calculation of Recycler Dynamic Aperture”, FERMILAB-TM-2143 (May 2001)
- [4] Etienne Forest, “Canonical Integrators as Tracking Codes (or How to Integrate Perturbation Theory with Tracking)”, SSC-138 (Sept. 1987)
- [5] J. A. MacLachlan, “Difference Equations for Longitudinal Motion in a Synchrotron”, Fermilab note FN-529 (1989)
- [6] V. I. Arnold, “Mathematical Methods of Classical Mechanics”, trans. K. Vogtmann and A. Weinstein, Springer Verlag, New York (1978)

Table 1: Definition of symbols in the difference equations

| Symbol | Meaning |
|-----------------|--|
| ϕ | rf phase |
| ϕ_s | synchronous phase |
| φ_i | difference between particle phase and ϕ_s |
| i | index for particles |
| n | index for turns |
| h | rf harmonic number |
| e | elementary particle charge (> 0) |
| E_s | synchronous energy |
| β | relativistic velocity v/c |
| γ | relativistic energy E_s/m_0c^2 |
| γ_T | γ of transition energy in synchrotron |
| V | total potential |
| ω_0 | angular frequency of beam circulation |
| ω_s | angular frequency of small amplitude synchrotron oscillations |
| ε_j | difference between energy of j^{th} particle and E_s |
| $S_{i,n}$ | phase slip/turn of i^{th} particle wrt synchronous particle |

Table 2: Parameters used in numerical tests of the synchrotron motion map

| Parameter | Symbol | Value | Units |
|---|-----------------|------------------------|-------|
| mean reference orbit radius | R_0 | 1000.0 | m |
| synchronous energy | E_s | | |
| above transition | | 150.938 | GeV/c |
| below transition | | 17.777 | GeV/c |
| transition energy/ m_0c^2 | γ_T | 18.6 | |
| slip factor $\gamma_T^{-2} - \gamma^{-2}$ | η | | |
| above transition | | $2.8515 \cdot 10^{-3}$ | |
| below transition | | $3.1249 \cdot 10^{-3}$ | |
| rf peak voltage | V_{rf} | 0.4 | MV |
| rf harmonic | h | 1113 | |
| synchrotron tune | ν_s | $1.157 \cdot 10^{-3}$ | |
| bucket height | H_B | 110 | MeV |

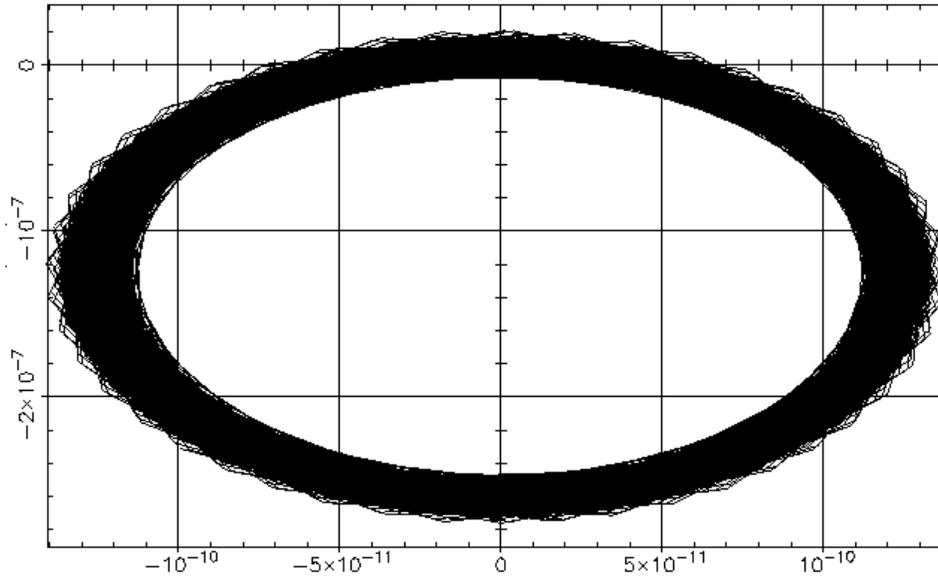


Figure 1: Locus of centroid of four macroparticles in a stationary Tevatron bucket at injection over 10^6 turns. The ordinate is in MeV difference from the synchronous energy and the abscissa is rf phase φ divided by the harmonic number $h = 1113$. The macroparticles closely surround the stable fixed point at $\Delta E = \pm 1$ MeV and $\Delta\varphi = \pm 1^\circ$.

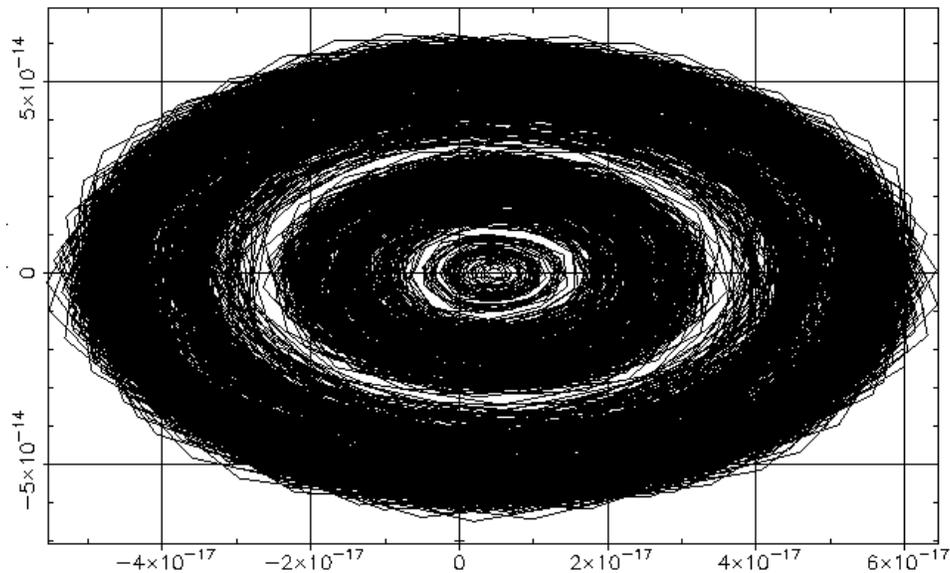


Figure 2: Similar to Fig. 1 except that the map is simplified to apply the same slip factor to all particles. The modified map is still symplectic, however.

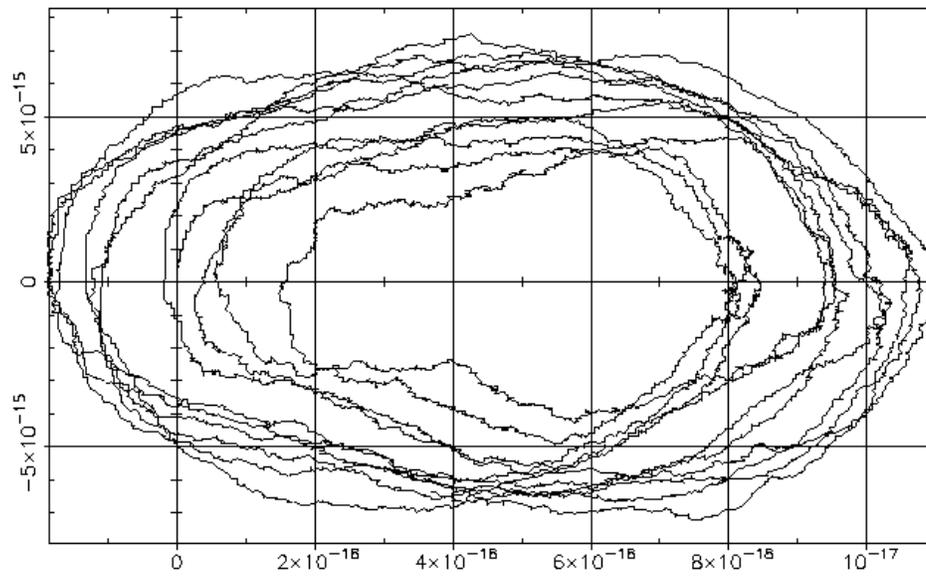


Figure 3: Similar to Fig. 2 except that only 10^4 turns are tracked

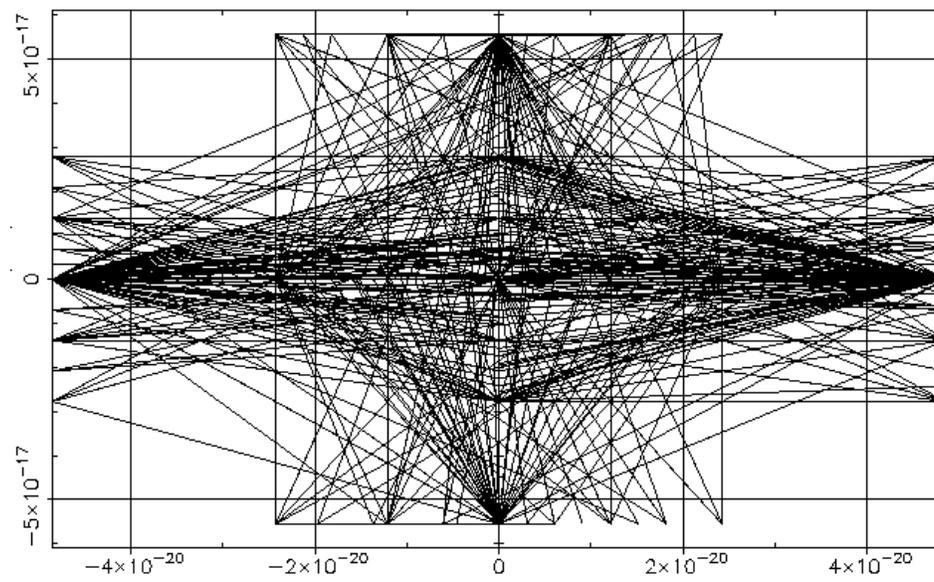


Figure 4: Similar to Fig. 3 except that the tracking is done at the energy below transition giving the same synchrotron tune and bucket area

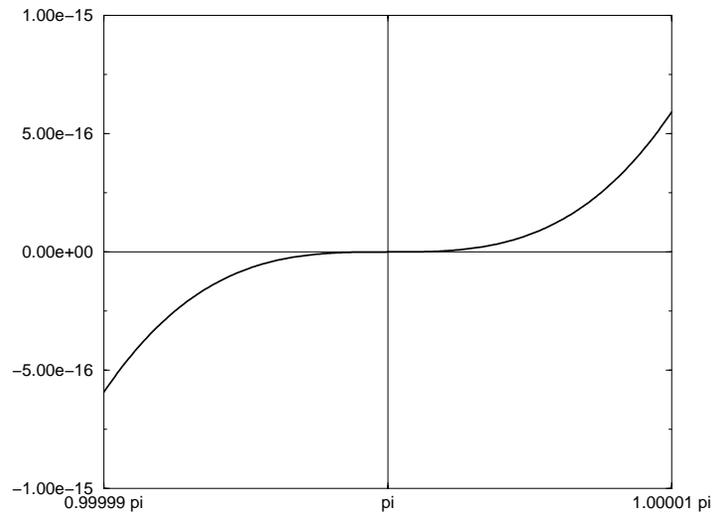


Figure 5: The third and higher order terms for the double precision sine function at $\pm 10^{-5}\pi$ around π , *i. e.*, the sine in this range less the difference between π and the argument

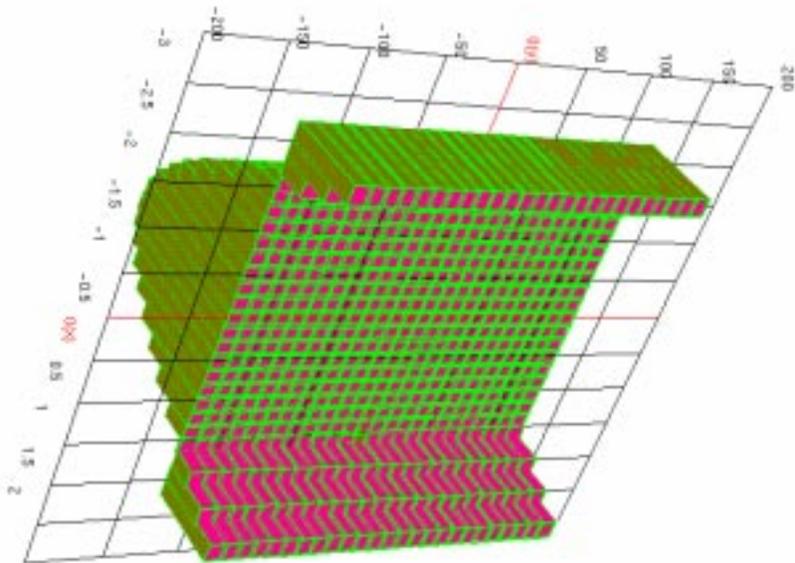


Figure 6: Plot of Jacobian determinant - 1 for synchrotron motion map on a grid with cells $0.2^\circ \times 0.2$ MeV; the parameters are those of a non-accelerating bucket at 150 GeV in the Tevatron accelerator.

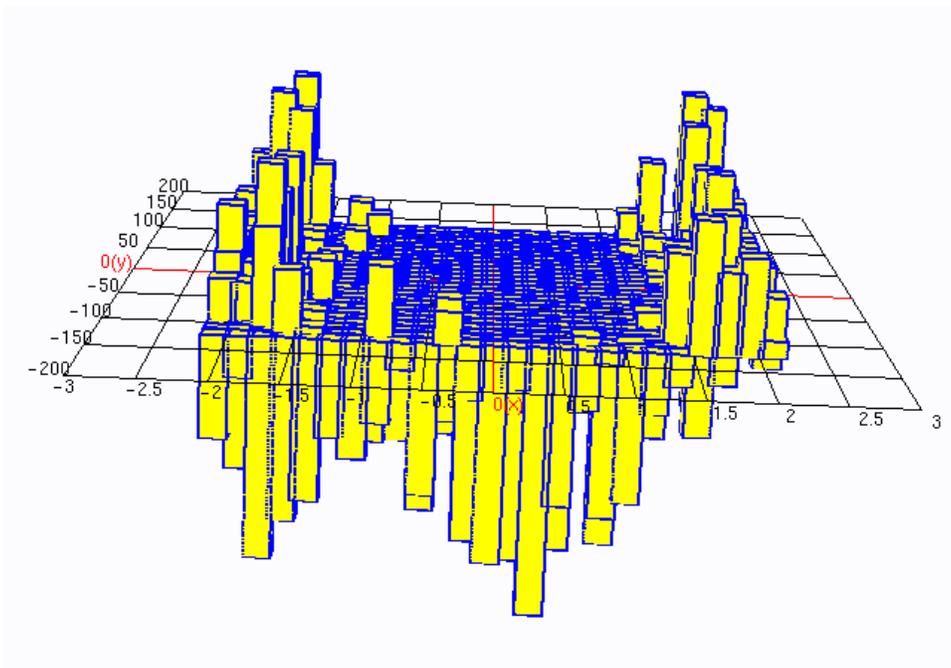


Figure 7: Like Fig. 6 with cell size $0.02^\circ \times 0.02 \text{ MeV}$

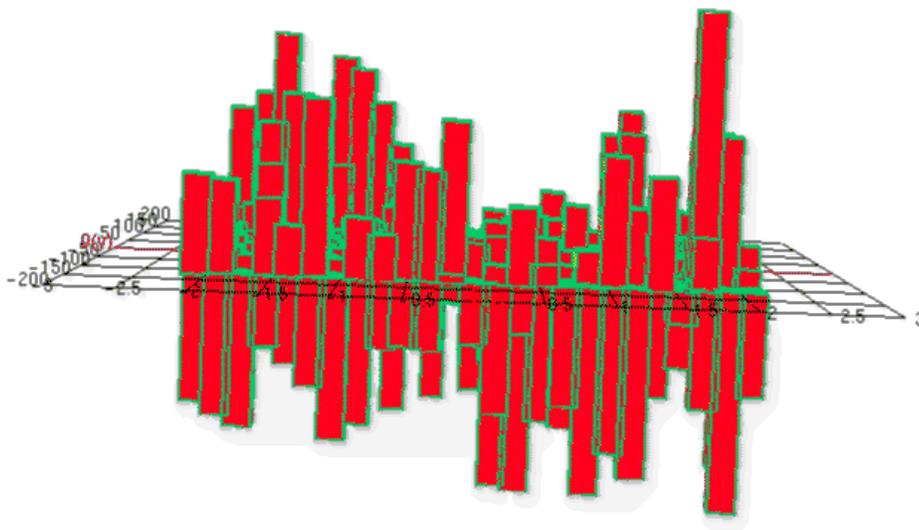


Figure 8: Like Fig. 6 with cell size $0.005^\circ \times 0.005 \text{ MeV}$

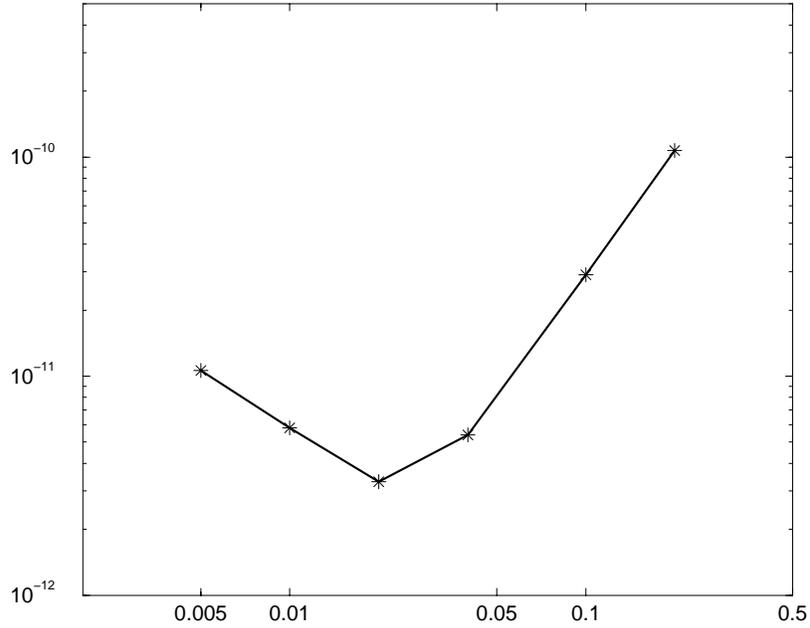


Figure 9: The absolute maximum Jacobian - 1 as a function of the side of the square cell used in evaluating the derivatives. The minimum in this curve establishes the cell size to find the best upper limits on the Jacobian -1 distribution.

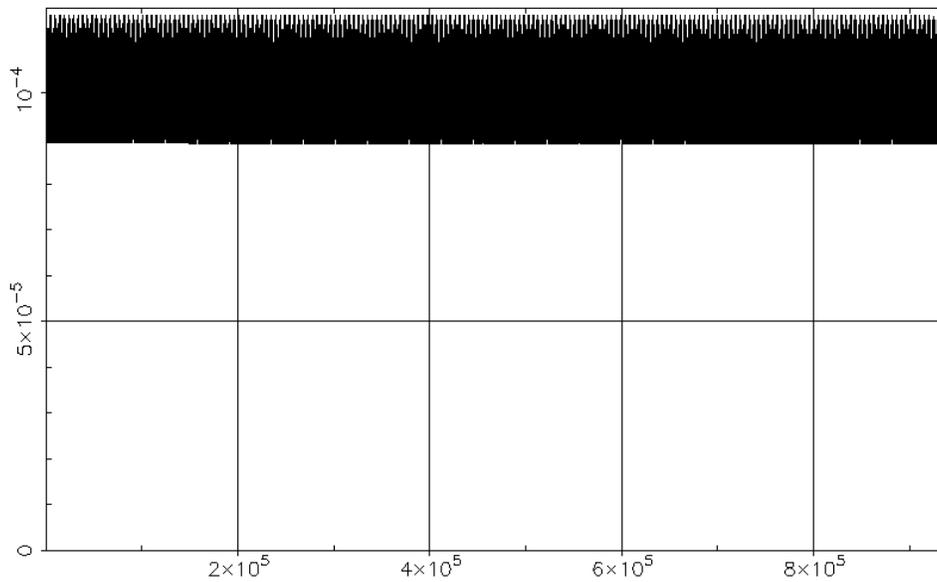


Figure 10: Fluctuation of the emittance during 10^6 turns for a nominal 0.5 eVs closed contour using Tevatron injection parameters. The ordinate is the difference between 0.5 and the rms emittance scaled to approximately 0.5.